

## Background error covariance estimation for atmospheric CO<sub>2</sub> data assimilation

Abhishek Chatterjee,<sup>1,2,3</sup> Richard J. Engelen,<sup>4</sup> Stephan R. Kawa,<sup>5</sup> Colm Sweeney,<sup>6,7</sup> and Anna M. Michalak<sup>2</sup>

Received 30 January 2013; revised 25 June 2013; accepted 16 July 2013; published 5 September 2013.

[1] In any data assimilation framework, the background error covariance statistics play the critical role of filtering the observed information and determining the quality of the analysis. For atmospheric CO<sub>2</sub> data assimilation, however, the background errors cannot be prescribed via traditional forecast or ensemble-based techniques as these fail to account for the uncertainties in the carbon emissions and uptake, or for the errors associated with the CO<sub>2</sub> transport model. We propose an approach where the differences between two modeled CO<sub>2</sub> concentration fields, based on different but plausible CO<sub>2</sub> flux distributions and atmospheric transport models, are used as a proxy for the statistics of the background errors. The resulting error statistics: (1) vary regionally and seasonally to better capture the uncertainty in the background CO<sub>2</sub> field, and (2) have a positive impact on the analysis estimates by allowing observations to adjust predictions over large areas. A state-of-the-art four-dimensional variational (4D-VAR) system developed at the European Centre for Medium-Range Weather Forecasts (ECMWF) is used to illustrate the impact of the proposed approach for characterizing background error statistics on atmospheric CO<sub>2</sub> concentration estimates. Observations from the Greenhouse gases Observing SATellite “IBUKI” (GOSAT) are assimilated into the ECMWF 4D-VAR system along with meteorological variables, using both the new error statistics and those based on a traditional forecast-based technique. Evaluation of the four-dimensional CO<sub>2</sub> fields against independent CO<sub>2</sub> observations confirms that the performance of the data assimilation system improves substantially in the summer, when significant variability and uncertainty in the fluxes are present.

**Citation:** Chatterjee, A., R. J. Engelen, S. R. Kawa, C. Sweeney, and A. M. Michalak (2013), Background error covariance estimation for atmospheric CO<sub>2</sub> data assimilation, *J. Geophys. Res. Atmos.*, 118, 10,140–10,154, doi:10.1002/jgrd.50654.

### 1. Introduction

[2] Atmospheric CO<sub>2</sub> observations provide a powerful constraint on net sources and sinks of CO<sub>2</sub>, as well as their spatial and temporal distribution [e.g., *Le Quéré et al.*, 2009]. The advent of several satellite-based instruments for observing CO<sub>2</sub> is expected to provide better insight into the critical controls over atmospheric CO<sub>2</sub> growth [e.g., *Scholes et al.*, 2009] and improve atmospheric flux inversions

[e.g., *Chevallier et al.*, 2009b; *Baker et al.*, 2010]. Assessing the information content of remote-sensing data sets, however, is difficult due to sampling limitations, data gaps, and incomplete error characterization in the observations. These data gaps can be caused by geophysical limitations, such as clouds and aerosols [e.g., *Bösch et al.*, 2006; *Tiwari et al.*, 2006], as well as by retrieval uncertainties. Further analysis in the form of statistical mapping [e.g., *Alkhaled et al.*, 2008; *Hammerling et al.*, 2012a, 2012b] or binning and averaging [e.g., *Tiwari et al.*, 2006; *Crevoisier et al.*, 2009; *Kulawik et al.*, 2010] is often necessary in order to create full-coverage maps and to reduce the errors associated with the data, thereby maximizing the usefulness of the data.

[3] An alternate approach is to use data assimilation (DA) techniques to extract information about global atmospheric CO<sub>2</sub> distributions from the available observations. Atmospheric CO<sub>2</sub> data assimilation [e.g., *Engelen et al.*, 2004; *Engelen and McNally*, 2005; *Engelen et al.*, 2009; *Liu et al.*, 2012] yields four-dimensional fields of atmospheric CO<sub>2</sub> concentrations that are statistically consistent with (1) the information provided by the CO<sub>2</sub> observations, and (2) additional sources of information such as model estimates of CO<sub>2</sub> fluxes, atmospheric transport, etc. These additional sources of information, commonly termed as the a priori model state (or the background state), are typically obtained from a short-range forecast valid at the

Additional supporting information may be found in the online version of this article.

<sup>1</sup>Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, Michigan, USA.

<sup>2</sup>Department of Global Ecology, Carnegie Institution for Science, Stanford, California, USA.

<sup>3</sup>Now at Data Assimilation Research Section, NCAR, Boulder, Colorado, USA.

<sup>4</sup>European Centre for Medium-Range Weather Forecasts, Reading, UK.

<sup>5</sup>NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

<sup>6</sup>Earth System Research Laboratory, NOAA, Boulder, Colorado, USA.

<sup>7</sup>Cooperative Institute for Research in Environmental Science, University of Colorado, Boulder, Colorado, USA.

Corresponding author: A. Chatterjee, Data Assimilation Research Section, NCAR Boulder, CO, USA. (abhishek@ucar.edu)

©2013. American Geophysical Union. All Rights Reserved.  
2169-897X/13/10.1002/jgrd.50654

time of assimilation. The background state incorporates knowledge about the mechanistic processes governing the carbon cycle and/or imposes physical or dynamical constraints on the assimilation. An attractive feature of the DA framework [e.g., Engelen *et al.*, 2009] is that, along with CO<sub>2</sub> observations, it is also possible to assimilate relevant meteorological variables such as temperature and humidity, which are known to affect the observed radiances from which CO<sub>2</sub> information is derived. The final analysis is a consistent estimate of the atmospheric CO<sub>2</sub> concentrations which, if produced with sufficiently high accuracy, can be used to improve estimates of CO<sub>2</sub> sources and sinks [e.g., Chevallier *et al.*, 2009a].

[4] In all DA applications, ranging from atmospheric CO<sub>2</sub> or other trace gases to Numerical Weather Prediction (NWP), the observations and the background are weighted [e.g., Nichols, 2010] based on the accuracy of the information sources. Like any other information source, the background state is prone to errors, which are accounted for through the so-called background error covariance statistics [e.g., Bannister, 2008a]. It is now well accepted that a realistic representation of the background error distribution is very important for successful data assimilation. The background error statistics serve to filter and spatially spread the information provided by the observations, and to impose correlation among different model variables. The critical role played by the background has been demonstrated by Cardinali *et al.* [2004], for example, where it was shown that only 15% of the information content of a well-balanced meteorological analysis is attributable to the assimilated observations; the remaining 85% of the information is provided by the background.

[5] A conceptual definition of background error is relatively trivial, in the sense that it corresponds to the statistical difference between the background atmospheric state and the true atmospheric state. Nonetheless, realistic estimation of the error statistics is not straightforward for several reasons. First, the true atmospheric state is never exactly known, such that the background errors and associated covariances must be estimated from surrogate data. A second difficulty is that error contributions to the background are relatively intricate, because the background is the result of a complex data assimilation procedure that involves interplay between the observations assimilated in the past, the analysis formulation, and a forecast model operator. A third difficulty is that the size of the full background error covariance matrix is too large to be stored explicitly, necessitating a variety of reduced-rank approximations to make the computations feasible.

[6] Various techniques have been developed to tackle these challenges [e.g., Bannister, 2008a, 2008b], and there is now a substantial literature on background error statistics for NWP, covering their nature, estimation, and practical implementation in operational settings [e.g., Derber and Bouttier, 1999; Fisher, 2004, 2006; Belo Pereira and Berre, 2006; Buehner and Charron, 2007; Pannekoek *et al.*, 2007, 2008; Raynaud *et al.*, 2009; Hess, 2010; Bonavita *et al.*, 2011; Brousseau *et al.*, 2012]. Similarly, there is an increasing focus on estimation of background error statistics for constituent assimilation applications [e.g., Benedetti and Fisher, 2007; Chai *et al.*, 2007; Constantinescu *et al.*, 2007; Kahnert, 2008; Singh *et al.*, 2011; Massart *et al.*, 2012], which requires a somewhat different approach than the NWP problem. This is primarily because time-varying boundary values are optimized instead of an initial condition (i.e., the atmospheric state) at the

start of a relatively short assimilation window. More recently, an increasing number of studies [e.g., Chai *et al.*, 2007; Buehner and Charron, 2007; Berre and Desroziers, 2010; Singh *et al.*, 2011] have demonstrated that, for both NWP and constituent assimilation applications, it is crucial to take into account the spatial correlations in the background errors, and therefore use nondiagonal error covariance matrices to represent these errors.

[7] Existing forecast or ensemble-based techniques, however, are problematic for atmospheric CO<sub>2</sub> data assimilation, as they are unable to explicitly take into account the significant uncertainty associated with the surface CO<sub>2</sub> fluxes. Spatial and temporal variations in the CO<sub>2</sub> fluxes are a key driver behind atmospheric CO<sub>2</sub> distributions, and consequently the measured CO<sub>2</sub> concentrations. Additionally, the errors associated with the CO<sub>2</sub> transport model are not adequately characterized in the majority of these approaches. The prescribed background error statistics are based only on the internal uncertainties in the transport model, which are not a realistic representation of the true transport model errors. Failure to capture the true magnitude of the transport errors and/or the uncertainty of the fluxes results in an underestimation of the background error statistics, which makes the assimilation devalue the observations in favor of the background. This underestimation can be severe, for example, during the Northern Hemisphere summer when the uncertainties in the background CO<sub>2</sub> fluxes are typically high. One solution to this problem is to artificially inflate the background error [e.g., Engelen *et al.*, 2009], where the magnitude of the inflation factor may be determined based on comparisons between the CO<sub>2</sub> model concentrations and independent surface observations. This strategy still does not account for the spatial error correlations due to the flux uncertainties, however.

[8] The primary goal of this paper is to outline and test an approach for parameterizing the background error covariance for atmospheric CO<sub>2</sub> data assimilation, in a way that includes the statistics of errors resulting from both flux and transport uncertainties. The proposed approach is based on the assumption that the difference between total CO<sub>2</sub> concentrations (henceforth, denoted as  $\Delta\text{CO}_2$ ) from two state-of-the-art global models is statistically representative of the background errors. Because any two models provide a limited sample of the true background error distribution, it is beneficial to use models that are different both in terms of the underlying fluxes and the transport fields driving them to capture realistic variability in the background error. The error statistics are then generated from the  $\Delta\text{CO}_2$  fields using spatial statistical tools.

[9] In this study, the resulting background error statistics are implemented within the atmospheric CO<sub>2</sub> 4D-VAR system at ECMWF. This system, as described in Engelen *et al.* [2009], is akin to a NWP-DA setup, in which CO<sub>2</sub> mixing ratios are constrained along with other atmospheric variables such as temperature, winds, surface pressure, and humidity, to obtain a consistent estimate of the atmospheric CO<sub>2</sub> concentrations. Experiments are designed to evaluate (a) whether the difference between two models can be used as a proxy for the statistics of the background errors, (b) the extent to which the representation of background errors has a discernible impact on CO<sub>2</sub> estimates, and (c) whether including more realistic statistics of errors improves the performance of the data assimilation system, relative to simply accounting for internal transport model uncertainties as available from a forecast-based

technique (standard version of the NMC method, according to the terminology of Široká *et al.* [2003]). The predictions of the 4D-VAR analyses are evaluated using independent observations of CO<sub>2</sub> from aircraft profiles and observations of column-averaged dry mole fractions of CO<sub>2</sub> (i.e., XCO<sub>2</sub>) from the Total Column Carbon Observing Network (TCCON).

[10] Although the application presented here primarily focuses on the operational ECMWF 4D-VAR system, the proposed  $\Delta$ -statistics approach is also applicable to ensemble data assimilation systems. Within an ensemble system, the  $\Delta$ -statistics approach may be used to define an initial background state, which would be refined as the background state of each ensemble member is perturbed further during the DA process. Additionally, this study focuses on atmospheric CO<sub>2</sub> data assimilation, but the  $\Delta$ -statistics approach is relevant for other trace gas assimilation applications, especially ones in which the background errors are influenced by both atmospheric transport and emission patterns.

## 2. Experimental Framework

### 2.1. Four-Dimensional Variational Data Assimilation

[11] The atmospheric 4D-VAR data assimilation system used in this study is based on the ECMWF Integrated Forecasting System (IFS) transport model with CO<sub>2</sub> fluxes prescribed at the surface based on climatological and inventory data [e.g., Engelen *et al.*, 2009; Hollingsworth *et al.*, 2008]. The system assimilates the same meteorological observations as the operational ECMWF system, along with observations constraining CO<sub>2</sub>.

[12] In addition to the specification of the background covariance matrix, which forms the primary focus of this work (section 2.2), the data assimilation system here includes two major changes relative to Engelen *et al.* [2009]. First, the IFS version has been updated, and the version used in this study is based on CY37r3 (ECMWF, IFS documentation CY37r2, 2011, <http://www.ecmwf.int/research/ifsdocs/CY37r2/index.html>), which became operational in November 2011 and incorporates several improvements to the atmospheric forecast model and the data assimilation system at ECMWF. Second, while the focus in Engelen *et al.* [2009] was on the assimilation of radiances from the Atmospheric Infrared Sounder (AIRS) [Aumann *et al.*, 2003] and the Infrared Atmospheric Sounding Interferometer (IASI) [Siméoni *et al.*, 1997], the current study uses Level 2 retrieval data from the Greenhouse Gases Observing Satellite “IBUKI” (GOSAT) [Kuze *et al.*, 2009; Yokota *et al.*, 2009], which are obtained from version 2.9 of the ACOS algorithm [e.g., O’Dell *et al.*, 2012; Crisp *et al.*, 2012]. Based on the recommendations of the ACOS team, only the high (H) gain observations with the master quality flag equal to “good” are used in the assimilation process, with the observation errors set to 1% of the observed value. Although first estimates for bias correction of the ACOS data do exist [Wunch *et al.*, 2011a], these are not used in the experiments in this study. The GOSAT data provide a stronger constraint on CO<sub>2</sub> near the surface relative to AIRS and IASI observations, but have relatively poor geographical coverage.

### 2.2. Specification of the Background Error Statistics

#### 2.2.1. NMC Method

[13] The background error covariance matrix used in Engelen *et al.* [2009] is based on the NMC (National Meteorological

Center, nowadays named National Center for Environmental Prediction) method (Parrish and Derber [1992]). The NMC method is based on the principle of using a surrogate quantity to represent the background errors, where the surrogate is typically chosen to be the differences between forecasts of different lengths valid at the same time. In its simplest form [e.g., Široká *et al.*, 2003], the NMC method is implemented by taking the differences between 24 h forecasts and 12 h forecasts over a 1 month period. The main advantage of this method is that the forecasts required for calculating the background error statistics are readily available during the DA process.

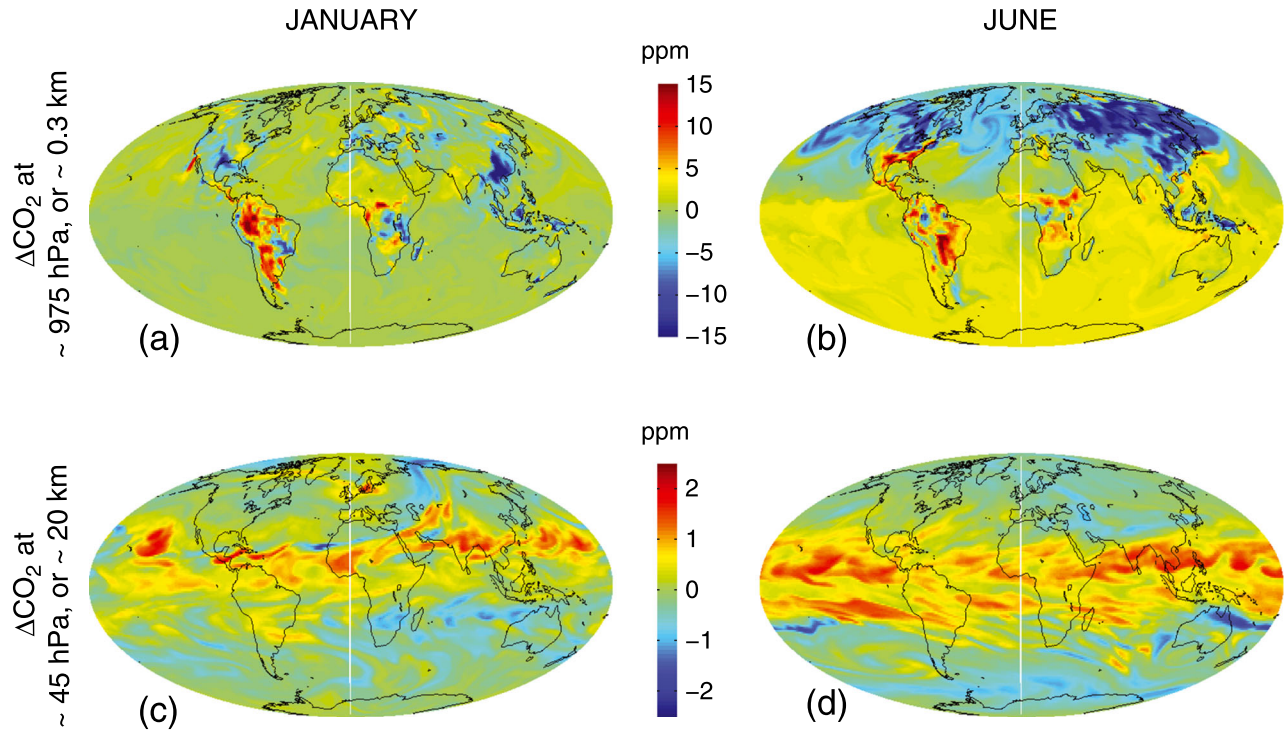
[14] One limitation common to all applications of the NMC method, is that the variance of background errors in data sparse regions is underestimated [e.g., Berre *et al.*, 2006] because differences between forecasts of different lengths are partially attributable to information from observations within the period between the starting times of the two forecasts. A second limitation, specifically for atmospheric CO<sub>2</sub> data assimilation, is that this approach does not account for the uncertainty associated with the underlying fluxes, because the two forecasts derive from the same set of prescribed fluxes. The NMC method is able to account for internal (i.e., within the same model) transport uncertainties only, and cannot represent systematic errors due to the transport model itself. The formalism of the NMC method (i.e., the analysis of increments) emphasizes the small-scale structures in the background CO<sub>2</sub>, missing much of the large-scale error patterns due to the dominant flux errors. Interestingly, this is different from NWP-related applications, where the observation errors are the main driver for the 6–12 h background forecast errors, and it has been pointed out [e.g., Berre *et al.*, 2006; Belo Pereira and Berre, 2006; Storto and Randriamampianina, 2010] that the NMC method likely overestimates the error correlations in that context. Finally, if there is no significant seasonality in the CO<sub>2</sub> observational constraint, then the error statistics defined via the NMC method remain invariant in time. Because the NMC statistics are only based on the internal (i.e., within model) transport uncertainties, they exhibit very little seasonal variability in the background errors.

[15] While there may be ways to extend the NMC method to take into account flux errors and seasonal variability in internal transport model uncertainties, these are neither straightforward to implement nor will account for the spatial correlation in the errors. Advances in the NMC method may also fail to take advantage of the fact that the forecasts can be obtained directly from existing runs; hence, the computational expense of generating the background error statistics may increase significantly.

#### 2.2.2. $\Delta$ -Statistics Method

[16] Analogous to the NMC method, the  $\Delta$ -statistics is formulated using a surrogate quantity, which is chosen to be the difference in the modeled CO<sub>2</sub> concentrations ( $\Delta$ CO<sub>2</sub>) from two global models. In this study, the two selected models are: (1) the GSFC Parameterized Chemistry and Transport Model (PCTM) [Kawa *et al.*, 2004] driven by analyzed meteorological fields from NASA’s Goddard Earth Observation System, version 4 (GEOS-4) with terrestrial biospheric sources and sinks based on computations of net primary productivity from the Carnegie-Ames Stanford Approach (CASA) [Randerson *et al.*, 1997] model and (2) the ECMWF Integrated Forecasting System (IFS) (<http://www.ecmwf>.





**Figure 1.**  $\Delta\text{CO}_2$  fields for (a and c) January and (b and d) June at two model levels. The fields are shown for the 15th day of the month at 1800 h UTC.

int/research/ifsdocs/CY37r2/index.html) model with the terrestrial biospheric fluxes prescribed from the Organizing Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) [Krinner *et al.*, 2005] model. Specifications for the other flux components (ocean, anthropogenic, and wildfire/biomass burning) are outlined in Table S1 in the supporting information. The main difference between the underlying fluxes is in the biospheric (CASA versus ORCHIDEE) and the anthropogenic (modeled versus compiled) CO<sub>2</sub> flux distributions, which, when propagated by different atmospheric transport (PCTM-GEOS4 versus IFS), result in two distinct specifications of CO<sub>2</sub> concentrations. Differences between the biospheric flux models have been reported in the literature [e.g., Xueref-Remy *et al.*, 2011; Huntzinger *et al.*, 2012], and we refer the reader to these studies for a more comprehensive overview.

[17] The differences between the two modeled CO<sub>2</sub> concentration fields ( $\Delta\text{CO}_2$ ; Figure 1) are obtained 3-hourly at a horizontal resolution of  $1^\circ \times 1.25^\circ$  for 60 sigma-hybrid levels of the atmosphere from which the horizontal and vertical error correlations are generated. Note that the two modeled CO<sub>2</sub> concentration fields are initialized using model-specific conditions, and these represent two independent long-term runs, which are not reinitialized during the analysis period. Global differences between the model runs are eliminated by subtracting any global offset between the models prior to the analysis. In this way, the error statistics used in the analysis represent long-term, rather than short-term, differences between the models and are likely to be conservative as they do not account for the reduction in background errors resulting from the assimilation of atmospheric data. By using conservative background error statistics, the approach allows the assimilation to put more weight on the observations.

[18] The vertical error correlations distribute the information of the observations in the vertical under the assumption that an error at a specific height is correlated with errors at other heights. The vertical error correlations are obtained by calculating the correlation coefficients of  $\Delta\text{CO}_2$  between different atmospheric (or model) levels. These vertical error correlations are constant in space but vary monthly to capture the seasonality in the  $\Delta\text{CO}_2$  field. The technique can be extended to consider different vertical error correlations for different geographic areas. Within a DA framework, both the vertical background error correlations and an averaging kernel constrain the model profile. In this study, only the vertical error correlations are obtained using the  $\Delta$ -statistics (or the NMC) method, while the averaging kernel is obtained from the GOSAT-ACOS data.

[19] The horizontal error correlations define how errors are correlated between grid boxes, and define the degree to which CO<sub>2</sub> is adjusted around grid boxes containing observations. The horizontal error correlations are obtained from a variogram analysis, a quantitative tool in geostatistics that has previously been used to characterize the spatial and temporal structure of atmospheric CO<sub>2</sub> [e.g., Alkhaled *et al.*, 2008; Hammerling *et al.*, 2012a]. The horizontal correlations are themselves spatially variable and are defined separately for each model level.

[20] Separate horizontal error statistics are calculated for each model level due to the difference in the patterns of CO<sub>2</sub> gradients at different levels of the atmosphere (Figure 1). Near the surface (Figures 1a and 1b), uncertainties in the prescribed fluxes result in higher errors in the background. In addition, the interaction between boundary layer dynamics and biospheric emissions (also known as “the rectifier effect,” Denning *et al.* [1999]) contributes to large variability in the

CO<sub>2</sub> concentrations near the surface. Conversely, in the free troposphere (Figures 1c and 1d), CO<sub>2</sub> is more dispersed and well-mixed, yielding smaller and smoother errors that are primarily impacted by the variability in atmospheric transport. Figure 1 shows the  $\Delta\text{CO}_2$  fields for two typical months—January and June 2010, which are representative of Northern Hemispheric winter and summer, respectively. As can be seen in this figure, the seasonality in the  $\Delta\text{CO}_2$  fields is more evident near the surface (Figures 1a and 1b) relative to higher levels in the atmosphere (Figures 1c and 1d).

[21] The plots in Figure 1 also show that within each model level, significant regional variability exists in the  $\Delta\text{CO}_2$  fields. These reflect the regional differences in surface fluxes between the models, as well as differences due to different representations of global atmospheric transport. Previous work by *Alkhaled et al.* [2008] and *Hammerling et al.* [2012a] using column-averaged CO<sub>2</sub> concentrations (i.e., XCO<sub>2</sub>) highlighted that significant regional variability exists in the spatial covariance structure of atmospheric CO<sub>2</sub> concentrations, and hence, any spatial analysis needs to be done regionally rather than globally. Keeping in mind that the  $\Delta\text{CO}_2$  field is a nonstationary process with variances and/or correlation lengths varying in space and time, we use a “moving window” analysis to calculate the horizontal error correlations.

[22] At any given model level, the regional variability of  $\Delta\text{CO}_2$  is quantified using the semivariogram ( $\gamma_{\text{region}}(d)$ ), as

$$\gamma_{\text{region}}(d) = \frac{1}{2} \left[ (\Delta\text{CO}_2(x_{\text{region}}) - \Delta\text{CO}_2(x_{\text{region}+d}))^2 \right] \quad (1)$$

where, similar to *Alkhaled et al.* [2008]: (a) regions are defined as overlapping 2000 km radius circles centered at each model grid cell, and (b)  $\gamma_{\text{region}}(d)$  is constructed using pairs of points, with the first point being within the specified region ( $\Delta\text{CO}_2(x_{\text{region}})$ ) and the other point being either within or outside that region ( $\Delta\text{CO}_2(x_{\text{region}+d})$ ). This approach accounts for both the observed variability within each subregion (by using all available pairs of points within a region) and large-scale variability (by using a random sample of the points outside the region). Once the regional variability of  $\Delta\text{CO}_2$  is defined, an exponential variogram model ( $\gamma_{\text{theo}}(d)$ ) is selected to represent the spatial autocorrelation structure. This model captures the decay in the spatial correlation as a function of separation distance ( $d$ ), parameterized by a variance ( $\sigma^2$ ) and a range ( $l$ ) parameter (equation (2)). While a variety of approaches has been suggested in the literature [e.g., *Chiles and Delfiner*, 2012] for fitting these parameters, we use the limited memory BFGS (L-BFGS) [*Nocedal and Wright*, 2006] minimization scheme for its computational efficiency. The horizontal error correlation coefficients are then obtained as the ratio of  $\mathbf{Q}(d)/\sigma^2$ , where

$$\begin{aligned} \mathbf{Q}(d) &= \sigma^2 - \gamma_{\text{theo}}(d) \\ \text{and, } \gamma_{\text{theo}}(d) &= \sigma^2 \left( 1 - \exp\left(-\frac{d}{l}\right) \right) \end{aligned} \quad (2)$$

[23] The choice of a simple isotropic exponential decay model is based on the nature of the underlying physical process and on the examination of the behavior of the variability, especially at smaller separation distances. Although

a more complex variogram model might be justified in areas with strong boundaries (e.g., near coasts), we err on the side of parsimony and keep the variogram model uniform while allowing its parameters to vary in space and time.

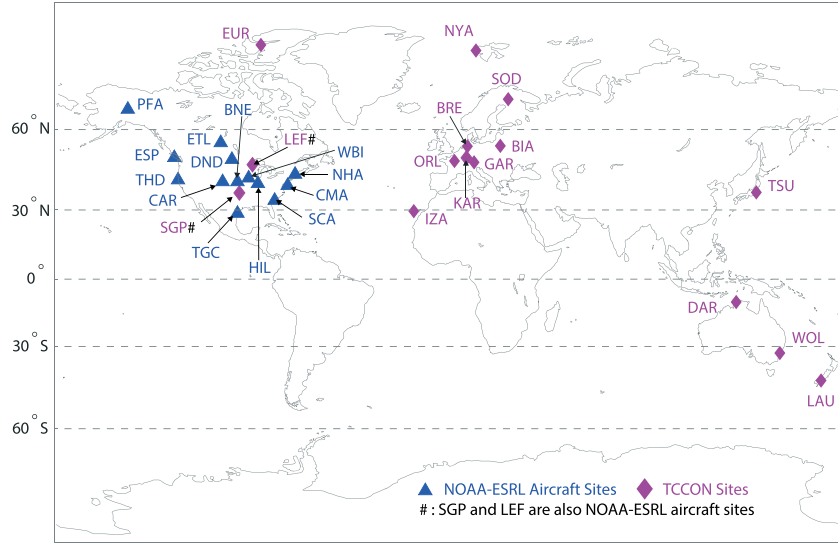
[24] The horizontal and vertical error correlations are generated for 1 day each month (on the 15th day at 1800 h UTC), and are assumed to be representative of the typical variability that would be observed during any individual day in the month. Ideally, the error correlations should be generated for all time periods (i.e., 3 hourly) within a month when the modeled CO<sub>2</sub> concentrations are available. This would avoid any form of spatial and/or temporal bias arising due to the way synoptic variability is handled in the individual models. At daily or subdaily time scales, however, any biases resulting from the selection of a particular set of models with their inherent physical processes, atmospheric transport patterns, etc., are mostly local and remain confined within the boundary layer. The error statistics did not vary substantially for different times of the day (for example, 0600 h UTC and 1800 h UTC) and for different days within a month. The decision to generate the error correlations once a month is thus primarily based on the tradeoff between reliably capturing monthly (or seasonal) variations in the background error and the computational time required to generate the error statistics. Ultimately, the appropriateness of this choice can be assessed by examining the fit between the assimilated CO<sub>2</sub> fields and independent data (section 4.3).

[25] The size of the background error covariance matrix is typically too large to be used directly within the ECMWF 4D-VAR system. Hence, the properties of the background error covariance inferred from the  $\Delta$ -statistics are subsequently modeled using the mathematical framework of wavelet-like, nonorthogonal basis functions having simultaneous localization, both in space and wave number. This mathematical formulation of the background error covariance matrix is based on *Fisher* [2004, 2006], and the application to tracer variables is also described in *Benedetti and Fisher* [2007].

### 3. Sample Application

#### 3.1. Experiments

[26] GOSAT observations for the year 2010 are assimilated into the 4D-VAR system to generate global 4D distributions of atmospheric CO<sub>2</sub>. Two independent sets of experiments are run, in which the background error covariance is prescribed based on the  $\Delta$ -statistics (henceforth “analysis with  $\Delta$ -statistics”) and the NMC approach (henceforth “analysis with NMC statistics”), respectively. While the background error using the  $\Delta$ -statistics method is defined exactly as outlined in section 2.2, the NMC method typically underestimates the background error. In order to offset this underestimation, following *Engelen et al.* [2009], the standard deviations of the background errors from the NMC method are inflated by a predetermined factor. The magnitude of the inflation factor is determined based on comparisons between the CO<sub>2</sub> model concentrations and independent surface and aircraft observations. All other parameters in these experiments are held the same to allow a straightforward evaluation of the impact on the CO<sub>2</sub> estimates due to different parameterization of the background error statistics.



**Figure 2.** Location of the aircraft and TCCON sites used in the study. The three-letter codes for the sites are defined in Table S2 in the supporting information.

[27] A third set of experiments is run in which fluxes are prescribed but no observational constraint on CO<sub>2</sub> is provided (henceforth “unconstrained model run”), to assess the impact of the GOSAT observations on the assimilation, and the impact of each of the background error representations. This experiment assimilates only meteorological observations and transports the CO<sub>2</sub> starting from the same initial field on January 1st 2010 as the other two sets of experiments.

### 3.2. Evaluation of the 4D CO<sub>2</sub> Fields From the Experiments

[28] Given that the true atmospheric CO<sub>2</sub> concentrations are unknown globally, the 4D-VAR estimates from the three experiments are evaluated using two sets of observations (Figure 2) that are not included in the assimilation process, namely column-averaged dry mole fractions of CO<sub>2</sub> (i.e., XCO<sub>2</sub>) from the Total Column Carbon Observing Network (TCCON) and vertical profiles of CO<sub>2</sub> from aircraft observations. Both of these observational data sets have much higher accuracy than the 4D-VAR analysis, and are assumed to be representative of the true atmospheric state.

[29] TCCON is a global network of calibrated ground-based Fourier transform spectrometers that measure the total column amount of various species (including CO<sub>2</sub>) by recording the direct solar spectra in the near-infrared spectral region [Wunch *et al.*, 2011b]. The TCCON observations are compared to cosampled column-averaged 4D-VAR analysis estimates within six latitudinal bands (90°–60°N, 60°–30°N, 30°–0°N, 0°–30°S, 30°–60°S, and 60°–90°S) to assess the impact of latitudinal differences in the background error statistics on the analyses. Because the GOSAT-ACOS XCO<sub>2</sub> data have already been calibrated using TCCON data, this comparison primarily evaluates the degree to which the  $\Delta$ -statistics approach improves the assimilation of GOSAT observations.

[30] Mean absolute errors (MAE) [e.g., Willmott and Matsuura, 2005] are calculated across all TCCON sites within each latitudinal band, as is the standard error of the mean:

$$\text{MAE} = \frac{1}{\sum_{i=1}^M N_i} \left[ \sum_{i=1}^M \sum_{j=1}^{N_i} |\hat{y}_{m_i, n_j} - y_{m_i, n_j}| \right] \quad (3)$$

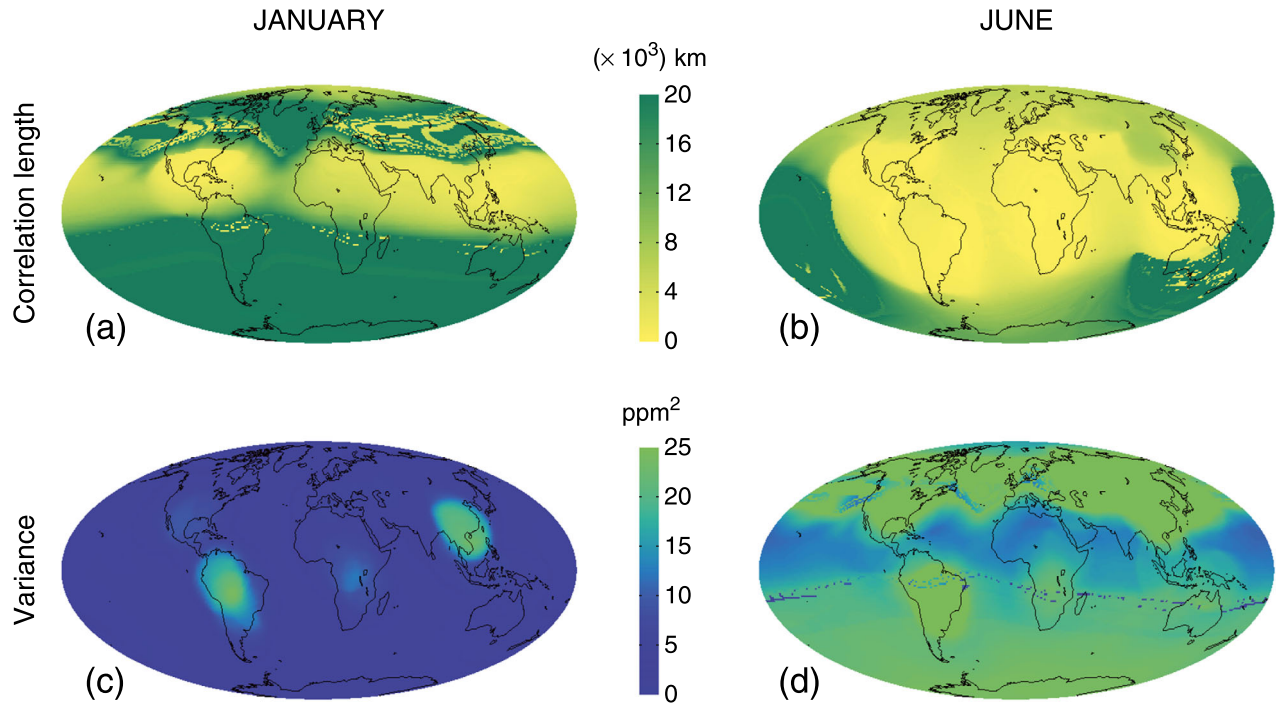
$$\sigma_{\text{MAE}} = \sqrt{\frac{1}{\left( \sum_{i=1}^M N_i \right)^2} \left[ \sum_{i=1}^M \sum_{j=1}^{N_i} |\hat{y}_{m_i, n_j} - y_{m_i, n_j}|^2 \right]} \quad (4)$$

where  $M$  represents the total number of TCCON sites,  $N_i$  represents the total number of observations at the  $i$ th TCCON site,  $y_{m_i, n_j}$  is the  $j$ th TCCON observation at the  $i$ th site, and  $\hat{y}_{m_i, n_j}$  is the corresponding 4D-VAR analysis value for that TCCON observation. Additionally, the MAE for each TCCON site is separately calculated as

$$\text{MAE}_i = \frac{1}{N_i} \left[ \sum_{j=1}^{N_i} |\hat{y}_{m_i, n_j} - y_{m_i, n_j}| \right] \quad (5)$$

[31] A second evaluation is carried out using aircraft observations available over North America. The National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA-ESRL) has been conducting long-term aircraft monitoring (<http://www.esrl.noaa.gov/gmd/ccgg/aircraft/>), in which vertical profiles of various trace gases (including CO<sub>2</sub>) are observed in the troposphere (i.e., surface to 8 km altitude) with high accuracy [e.g., Tans, 1996; Crevoisier *et al.*, 2010]. Sampling frequencies are weekly or biweekly for most sites. The aircraft observations and the sampled 4D-VAR analysis estimates are divided into four altitude bins, 0–2, 2–4, 4–6, and 6–8 km. For each altitude bin, equations equivalent to equations (3)–(5) are applied to calculate the MAE metrics between the 4D-VAR analysis estimates and the aircraft observations. This evaluation exercise provides quantitative information regarding the impact of the background error statistics in the vertical direction. Because the aircraft data have not been used to calibrate the GOSAT data and were not assimilated, this comparison primarily evaluates whether the  $\Delta$ -statistics





**Figure 3.** Inferred  $\Delta\text{CO}_2$  correlation length ( $3l$ ) and variance ( $\sigma^2$ ) parameters for (a and c) January and (b and d) June. The covariance parameters are shown for the 15th day of the month, 1800 h UTC at 975 hPa (i.e.,  $\sim 0.3$  km above the surface). Note the shorter correlation length and higher variance for June, indicative of greater variability in the surface CO<sub>2</sub> processes during this month.

approach yields more realistic atmospheric CO<sub>2</sub> fields. Finally, because aircraft observations are available only over North America for the study period, only a single latitude band from 90° to 0°N is considered for the aircraft evaluations.

[32] As the GOSAT XCO<sub>2</sub> retrievals mature and atmospheric CO<sub>2</sub> fields are obtained for more years, similar evaluation exercises using a variety of aircraft observations (e.g., CARIBIC—Civil Aircraft for the Regular Investigation of the Atmosphere Based on an Instrument Container, [Brenninkmeijer *et al.*, 2007]; CONTRAIL—Comprehensive Observation Network for Trace gases by Airliner, [Machida *et al.*, 2008]; HIPPO—HIAPER Pole-to-Pole Observations, [Wofsy, 2011]) are planned. These data sets have the additional advantage that they are available beyond the North American domain. Some of them also provide upper tropospheric data, which will allow for rigorous evaluation with these independent data at higher levels of the atmosphere.

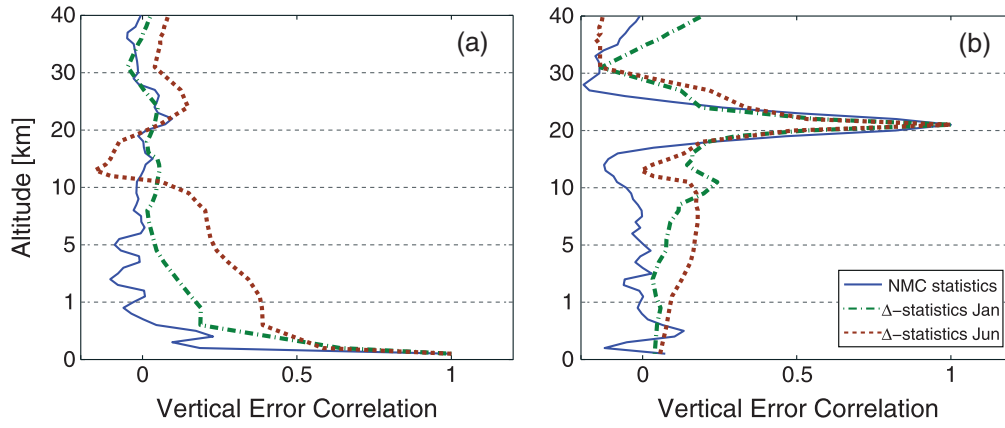
## 4. Results

### 4.1. Background Errors From the $\Delta$ -Statistics Approach

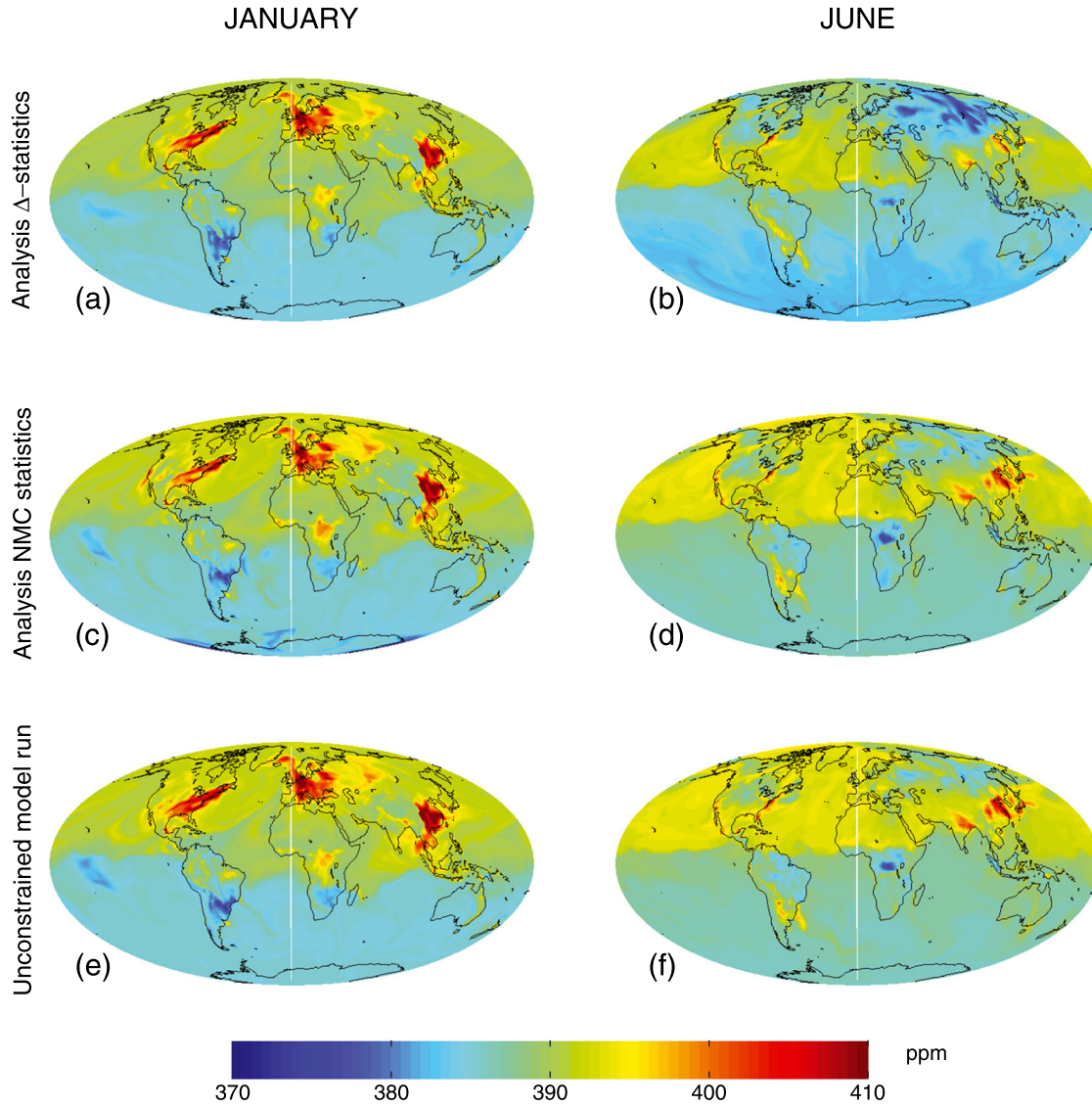
[33] Figure 3 shows the horizontal error correlation length and the variance for the near-surface  $\Delta\text{CO}_2$  fields shown in Figures 1a and 1b. The covariance parameters capture both regional and seasonal differences as seen in the  $\Delta\text{CO}_2$  fields. In June, strong biospheric signals dominate the Northern Hemisphere flux uncertainty, and variability is generally high relative to January. Consequently, the horizontal error correlations in Figures 3b and 3d show higher variance and shorter correlation lengths over the Northern Hemisphere relative to those presented in Figures 3a and 3c. The spatial patterns detected by the  $\Delta$ -statistics include large regions with

relatively low variances of below 2.5 ppm<sup>2</sup> during the Northern Hemisphere winter (January) and high variances reaching a maximum of 25 ppm<sup>2</sup> during the Northern Hemisphere summer (June). In June, regions with highly variable surface fluxes (for example, over Boreal forests—50° to 70°N, Figures 3b and 3d) also correspond to regions with large differences between models, and yield high estimated background errors, with variances of up to 25 ppm<sup>2</sup> and short correlation lengths ( $3l$ ) below 5000 km. Expectedly, analyses at higher model levels (Figure S1 in supporting information) show longer correlation lengths and lower variances relative to the surface. For example, over the Boreal forests, the variances reduce to 0.3 ppm<sup>2</sup> with long correlation lengths ( $3l$ ) of 15,000 km at 20.0 km elevation. Higher in the atmosphere, the impact of the near-surface CO<sub>2</sub> exchange is dampened, and the variability is dominated exclusively by the synoptic-scale mixing of air masses.

[34] During the Northern Hemisphere winter, and at higher levels of the atmosphere, differences between the  $\Delta\text{CO}_2$  fields are more affected by differences in the atmospheric transport rather than the uncertainty in the surface fluxes. This is reflected in the correlation length plots (Figure 3a and Figure S1a and S1b in the supporting information) where “streaks” with especially high or low correlation lengths are evident, corresponding to atmospheric transport pathway differences between the models [e.g., Stohl *et al.*, 2002]. These patterns in the  $\Delta\text{CO}_2$  field may also be a product of the different ways in which synoptic systems, for example, convection and turbulence associated with frontal activity, are being handled by the two transport models. In a separate set of tests, the horizontal error correlations were calculated for  $\Delta\text{CO}_2$  fields from different days and different times.

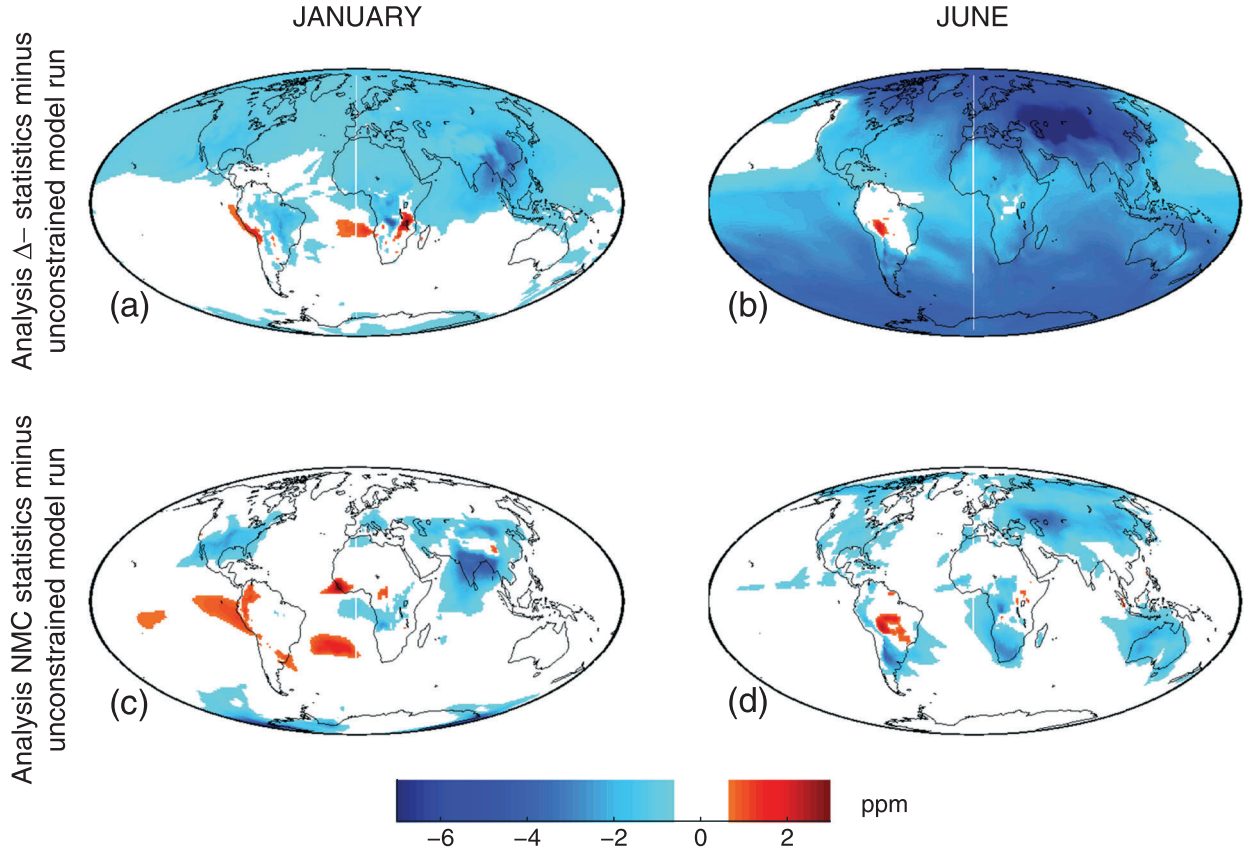


**Figure 4.** Vertical error correlations from the NMC method (constant throughout the year) and the  $\Delta$ -statistics (for January and June) at (a) the surface, and at (b)  $\sim 45$  hPa (i.e.,  $\sim 20$  km above the surface).



**Figure 5.** CO<sub>2</sub> concentrations from the 4D-VAR analyses based on the two background error statistics — (a and b)  $\Delta$ -statistics, (c and d) NMC statistics, and (e and f) the unconstrained model run for January and June 2010. The plots are shown for the 15th day of the month, 1800 h UTC at 975 hPa (i.e.,  $\sim 0.3$  km above the surface).



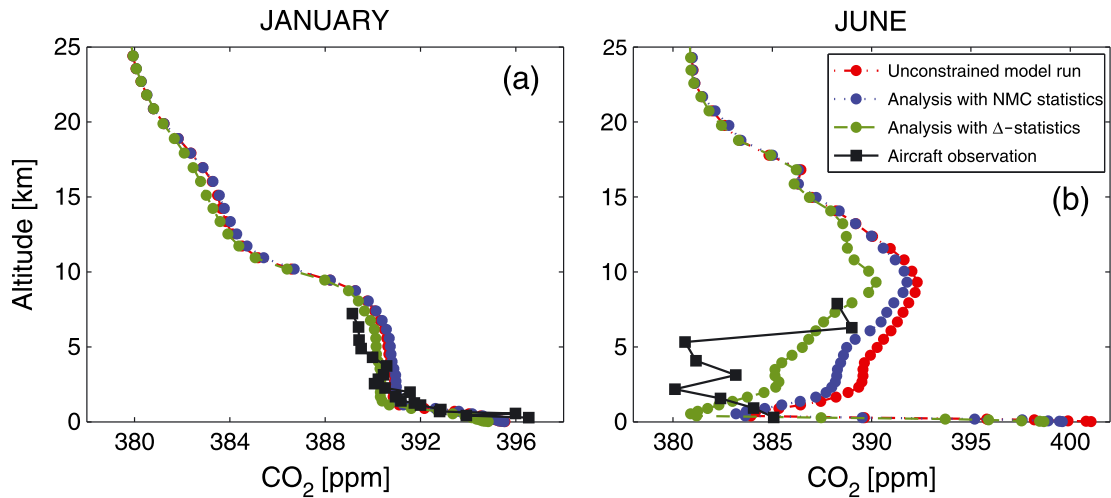


**Figure 6.** Monthly-averaged differences between the 4D-VAR analyses based on the two background error statistics, and the unconstrained model run, at 975 hPa (i.e.,  $\sim 0.3$  km above the surface) for (a and c) January and (b and d) June. Figures 6a and 6b represent the difference between the concentrations obtained from the analysis with  $\Delta$ -statistics and the unconstrained model run, and Figures 6c and 6d represent the difference between the concentrations obtained from the analysis with NMC statistics and the unconstrained model run.

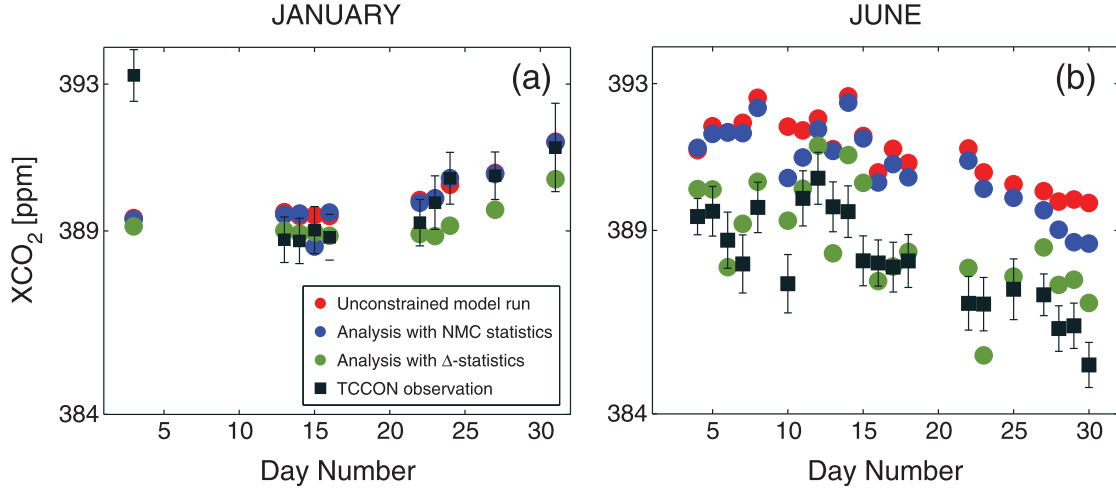
The overall structure and pattern of the horizontal error correlation values were found to be similar (results not shown) to those reported in Figure 3 (and Figure S1); hence, the patterns in the horizontal error correlations can be assumed

to be representative of the variability in the atmospheric transport rather than artifacts due to synoptic variability.

[35] The vertical error correlations (Figure 4) based on the  $\Delta$ CO<sub>2</sub> fields show distinct seasonality and decay more



**Figure 7.** Evaluation of 4D fields of CO<sub>2</sub> using aircraft observations over Worcester, Massachusetts (site code – NHA) on (a) January 9th, 2010, and (b) June 26th, 2010. CO<sub>2</sub> profiles are shown only up to 25 km to show the differences among the 4D-VAR analyses near the surface.

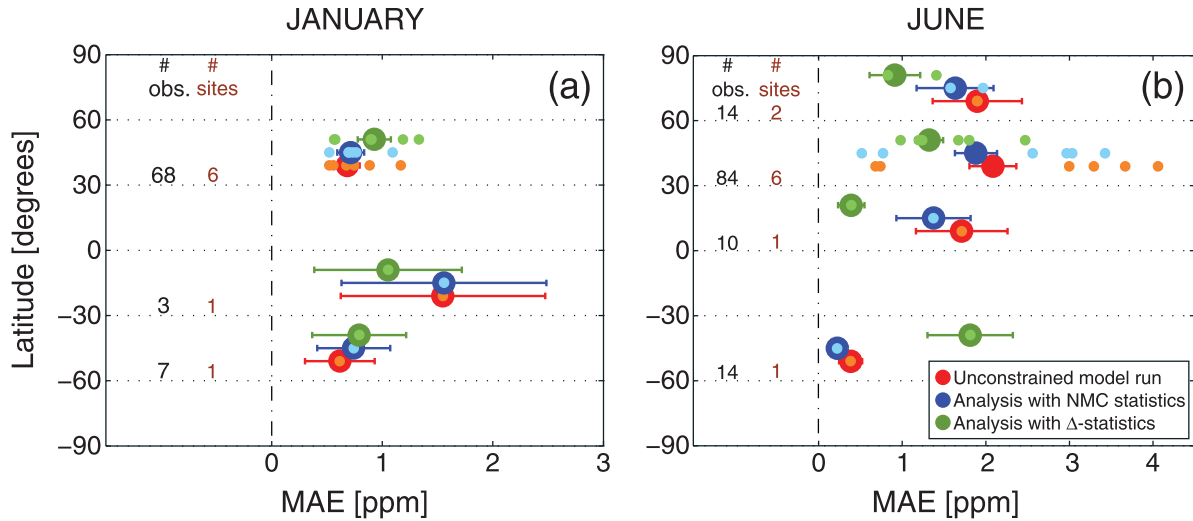


**Figure 8.** Evaluation of column-averaged CO<sub>2</sub> using TCCON observations over Bialystok, Poland (site code – BIA) during (a) January 2010, and (b) June 2010. The error bars represent the 95% uncertainty on the TCCON observations.

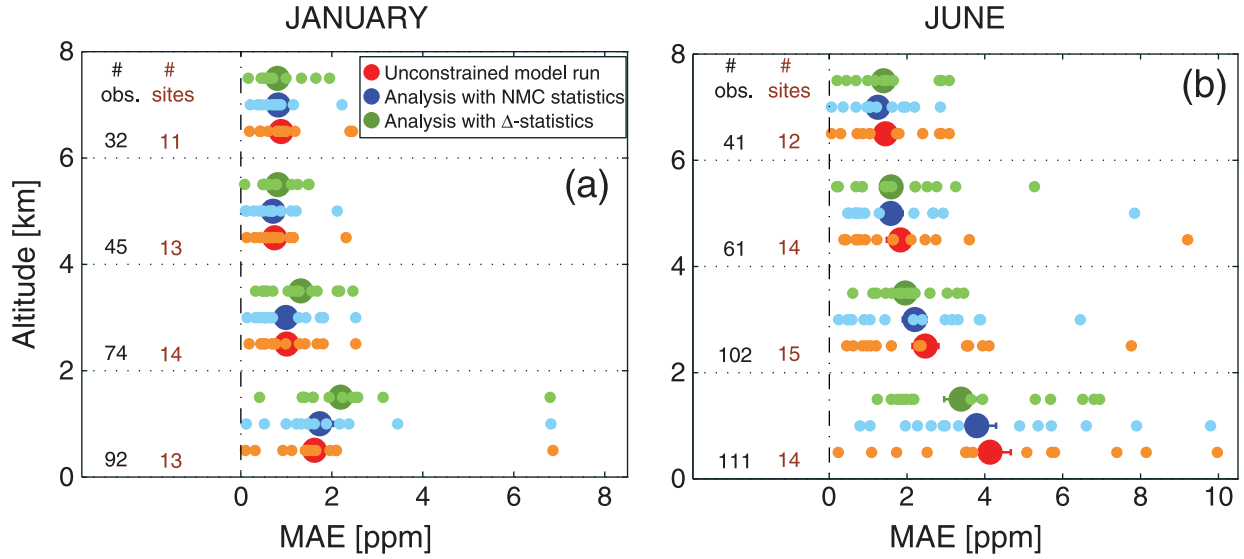
gradually across model levels (Figures 4a and 4b—between 10 and 0 km) relative to those derived from the NMC method. This gradual decay implies that the analysis will spread the information from local observations more in the vertical relative to the analysis using the  $\Delta$ -statistics. Generally speaking, the vertical error correlations from the  $\Delta$ -statistics by design tend to capture features associated with the uncertainty in atmospheric transport better than the error characteristics using the NMC statistics. During the month of June, for example, the error correlations between the surface and higher levels derived from the  $\Delta$ -statistics drop sharply to negative values around 10–12 km, which coincides with the location of the tropopause. While the exact causes for these

negative correlations remain unclear, reasons related to the uncertainty in the atmospheric transport can be hypothesized.

[36] First, examination of the CO<sub>2</sub> fields shows strong negative gradients at these atmospheric levels (~10–12 km), especially over the extra-tropical regions. The role of the tropopause as a barrier to isentropic mixing is well-known [e.g., Holton *et al.*, 1995; Andrews *et al.*, 2001; Bönisch *et al.*, 2009], and most likely causes negative gradients in the CO<sub>2</sub> concentrations to develop, while also damping the amplitude of the seasonal cycle of CO<sub>2</sub> [e.g., Gurk *et al.*, 2008]. When the gradients or amplitude of the CO<sub>2</sub> concentrations near the tropopause are not simulated similarly between the two models, large errors (i.e., differences in model)



**Figure 9.** MAE between daily column-averaged CO<sub>2</sub> estimates from the 4D-VAR analyses and the TCCON observations, binned by latitude for (a) January 2010, and (b) June 2010. MAE between the 4D-VAR analyses and observations from each TCCON site are shown as the smaller dots. The error bars on the MAE represent the standard error of the mean ( $\pm \sigma_{\text{MAE}}$ , equation (4)) and take into account the total number of independent observations (# obs. as reported in black text), available within each latitude band. Values reported in brown denote the number of TCCON sites (# sites), available within each latitude band.



**Figure 10.** MAE between the CO<sub>2</sub> estimates from the 4D-VAR analyses and the aircraft CO<sub>2</sub> observations, binned by altitude for (a) January 2010, and (b) June 2010. MAE between the 4D-VAR analyses and observations from each aircraft site are shown as the smaller dots. The error bars on the MAE represent the standard error of the mean ( $\pm \sigma_{MAE}$ , equation (4)) and take into account the total number of independent observations (# obs. as reported in black text) available within each altitude band. Values reported in brown denote the number of aircraft sites (# sites) available within each altitude band.

are inferred relative to the other atmospheric levels. Second, if the two transport models predict different tropopause heights, this may combine with the steep gradients at this level and result in relatively large differences. If these differences are systematic globally, it could create negative correlations with the surface, which is coupled more tightly with the upper troposphere than the lower stratosphere. Finally, the tropopause errors could be anti-correlated with errors at the surface as a result of local convective vertical transport (pumping) differing between the two transport models. Beyond the tropopause, however, large-scale meridional stratospheric circulation dominates, which is simulated similarly by both the models, resulting in a gradual decrease in the inferred errors.

[37] The error statistics also provide an initial indication of the observational impact on the analysis. The long  $\Delta$ CO<sub>2</sub> correlation lengths are a result of the large-scale errors in the CO<sub>2</sub> background, which primarily arise from the large-scale errors in the flux fields. Within a DA analysis, specifying these long error correlations will induce the analysis to suppress small-scale features while propagating the information from the observations to long distances, in the horizontal and the vertical. If the flux fields differed only on smaller scales, for example, if the signal were dominated by fossil fuel emissions, shorter correlation lengths would have been derived. Thus, one advantage of the proposed  $\Delta$ -statistics method lies in inferring error correlations that directly depend on the uncertainty associated with all the underlying drivers of the physical problem being examined.

[38] If the  $\Delta$ CO<sub>2</sub> fields are a reasonable representation of the structure of the background errors, then the corresponding horizontal and vertical error correlations indicate that the background error statistics should vary spatially and temporally. By specifying these error statistics to be constant over large areas or forcing them to be invariant in time, the NMC statistics

may underestimate (or over-estimate) the magnitude of the true errors and therefore degrade the DA analysis.

#### 4.2. Impact of Background Error Statistics on 4D-VAR Analysis

[39] Figure 5 shows a sample of atmospheric CO<sub>2</sub> fields near the surface from the three 4D-VAR experiments (see Figure S2 in supporting information for a sample of atmospheric CO<sub>2</sub> fields at ~45 hPa, or ~20 km elevation above the surface). All three experiments capture the latitudinal and interhemispheric gradients in CO<sub>2</sub> as well as the seasonal cycle of the Northern Hemisphere biospheric CO<sub>2</sub>. Evaluation of the two 4D-VAR analyses using the GOSAT observations against the unconstrained run indicates that the analysis using the  $\Delta$ -statistics allows more pronounced adjustments by the GOSAT observations than the analysis using the NMC statistics. A typical example is visible over boreal Asia, where the CO<sub>2</sub> fields in the analysis with  $\Delta$ -statistics (Figure 5b) are more different from the unconstrained model run (Figure 5f) than the CO<sub>2</sub> fields in the analysis with NMC statistics (Figure 5d) from the unconstrained model run (Figure 5f).

[40] The  $\Delta$ -statistics and the NMC statistics specify different changes (in terms of both pattern and magnitude) to the background model CO<sub>2</sub> concentrations. This can be examined by looking at the monthly-averaged differences between the 4D-VAR analyses based on the two background error statistics and the unconstrained model run (Figure 6) or by looking at the CO<sub>2</sub> analysis increments (analysis minus background values), which show the direct impact of the observations on the CO<sub>2</sub> field (Figure S3 in supporting information). The CO<sub>2</sub> increments in the NMC experiment are generally small scale, while the  $\Delta$ -statistics allow much broader scale adjustments to the CO<sub>2</sub> fields. This supports the original hypothesis that the longer error correlation lengths (both in the horizontal and the vertical) in the  $\Delta$ -statistics allow the



GOSAT observations to have a greater impact (Figures 6a and 6b), and make adjustments over larger scales. The smaller error variances associated with the NMC statistics give less weight to the GOSAT observations in favor of keeping the analysis close to the background CO<sub>2</sub> distribution and produce only localized changes (Figures 6c and 6d). The mean absolute difference at 975 hPa (i.e., ~0.3 km above the surface) in CO<sub>2</sub> concentrations between the analysis with the NMC statistics and the unconstrained model run (January—0.22 ppm, June—0.36 ppm) is smaller, and less variable across seasons, than the corresponding difference between the analysis with the  $\Delta$ -statistics and the unconstrained model run (January—0.63 ppm; June—3.3 ppm).

[41] The difference between the two analyses is greater near the surface, where most of the variability in the CO<sub>2</sub> processes (such as fossil fuel emissions, biospheric exchanges, influence of boundary layer variations, etc.) is present. This can be primarily attributed to the fact that at lower levels of the atmosphere, (a) the error statistics from the NMC and the  $\Delta$ -statistics methods are most different and (b) the GOSAT observations tend to have the largest impact on the analysis. The accuracy of these adjustments is next evaluated using independent observations.

### 4.3. Evaluation of 4D-VAR Analysis Using Independent Observations

[42] The impact of the background error statistics on the 4D-VAR analysis, in terms of yielding a more realistic atmospheric CO<sub>2</sub> product and a better assimilation system, is assessed by comparing the estimated atmospheric CO<sub>2</sub> fields to observations from the aircraft and the TCCON networks, respectively. Figure 7 shows example of two profiles collected on January 9th, 2010 and June 26th, 2010, over Worcester, Massachusetts (site code-NHA in Figure 2). In both cases, the 4D-VAR analyses reduce the mismatch between the background and the true atmospheric CO<sub>2</sub> state. The ability of the GOSAT observations to adjust the details of the background CO<sub>2</sub> profile, however, is limited by the NMC method relative to the more flexible  $\Delta$ -statistics approach. Figure 8 shows the TCCON observations collected over Bialystok, Poland (site code—BIA in Figure 2) for different days in the months of January and June 2010. The NMC statistics again limit the degree to which the GOSAT observations can adjust the background state.

[43] Overall, the CO<sub>2</sub> estimates based on the  $\Delta$ -statistics show significant improvement relative to the analysis using NMC statistics during the Northern Hemisphere summer, when significant uncertainty in the fluxes is present (Figure 9). The variability of the MAE across TCCON sites is also smaller when using the  $\Delta$ -statistics approach. This indicates that the CO<sub>2</sub> concentrations obtained from the analysis with the  $\Delta$ -statistics are consistently closer to the observed CO<sub>2</sub> from the TCCON sites. As outlined earlier, the GOSAT data are themselves calibrated to agree with the TCCON XCO<sub>2</sub> observations; hence, better agreement of the analysis represents an unambiguous improvement. During the Northern Hemisphere winter, the analysis with the  $\Delta$ -statistics performs similarly to the analysis using NMC statistics (Figure 9a, 30°–60°N). This is not surprising given the dearth of good-quality GOSAT retrievals, as a consequence of which the impact of observations on the analysis is not significant. Similarly, the performance over the Southern Hemisphere winter

(Figure 9b, 30°–60°S) is difficult to judge given the paucity of TCCON sites and wintertime CO<sub>2</sub> data.

[44] The aircraft profiles allow us to also examine the assimilation performance at different levels of the atmosphere. As seen from Figure 7, at higher levels of the atmosphere (~ ≤ 45 hPa or ≥ 20 km), the differences in the analysis estimates become negligible due to the following: (a) the background error covariance from the  $\Delta$ -statistics and the NMC statistics being similar and (b) the GOSAT observations being less informative in the free troposphere and stratosphere relative to the lower troposphere. Between 10 and 20 km (i.e., ~200 hPa and 45 hPa) only occasional differences are visible between the 4D-VAR analyses and the unconstrained model run, but both analyses remain close to the background model CO<sub>2</sub>.

[45] These patterns are not limited to the particular aircraft profile examined in Figure 7 but are also true for other aircraft profiles (Figure 10). Given that most of the variability is driven by changes at the surface, information from the observations can only change CO<sub>2</sub> concentrations at upper levels of the atmosphere through atmospheric transport or through the vertical error correlations prescribed in the background covariance matrix. As discussed earlier in section 4.1, the  $\Delta$ -statistics is expected to spread information more strongly in the vertical relative to the NMC statistics. Whereas the MAE for the analyses with the  $\Delta$ -statistics and the NMC statistics are similarly higher up in the atmosphere (Figures 10a and 10b, 6–8 km), differences in the MAE are more clearly visible lower down (Figures 10a and 10b, 0–2 and 2–4 km). Similar to the TCCON evaluation, the analysis with the  $\Delta$ -statistics outperforms the analysis using the NMC statistics and the unconstrained run slightly during the Northern Hemisphere summer (Figure 10b, 0–2 km), but does slightly worse (Figure 10a, 0–2 km) during the winter. During the winter months, even the analysis with the NMC statistics does worse than the unconstrained model run (Figure 9a—30°–60°N and Figure 10a—0–2 km), which points to an inconsistency between the constraint provided by the GOSAT data and the available independent observations from the aircraft samples.

[46] The 4D-VAR configuration assumes unbiased Gaussian error statistics. Hence, any systematic differences between the observations and the model background may be interpreted incorrectly as an uncertainty rather than a bias. When the background error is small relative to the observation error, as is the case with the NMC experiment, the analysis will be directed toward the model background, irrespective of the bias in either the observations or the model background itself. When the background error is large compared to the observation error, as is the case in the  $\Delta$ -statistics experiment, the analysis is pulled toward the observations in the areas where they are available. Ultimately, this leads to the observed differences in the final CO<sub>2</sub> analysis. Previous experiments (results not shown) with a biased set of AIRS data demonstrated that the observational constraint cannot be biased for the  $\Delta$ -statistics to provide reasonable results. The observed poorer performance of the  $\Delta$ -statistics during the winter months could therefore also be a by-product of biases in the GOSAT XCO<sub>2</sub> retrievals. In this case, erroneous data with loose background error statistics (i.e.,  $\Delta$ -statistics) are potentially more harmful than with a more constrained background error statistics (i.e., NMC statistics). The challenge, as in all DA systems, remains to achieve the

best balance (i.e., weighting) between the observation and the background information sources.

## 5. Conclusions

[47] The specification of background error covariance statistics is a critical component of any atmospheric data assimilation system, and accurate representation of these statistics is necessary in order to make efficient use of the observational information. This study examined the specification of the background error statistics for the atmospheric CO<sub>2</sub> data assimilation problem. Using the state-of-the-art 4D-VAR system in place at ECMWF, this study shows that it is necessary to modify the forecast or ensemble-based techniques prevalent in NWP applications to account for both the errors in the underlying CO<sub>2</sub> fluxes and the errors associated with atmospheric transport of CO<sub>2</sub>. Limitations associated with existing methods, such as the NMC method, prompted the investigation of a new flexible approach for parameterizing the background error statistics that is more suited to atmospheric CO<sub>2</sub> data assimilation.

[48] Using the difference between CO<sub>2</sub> concentrations resulting from alternate sets of CO<sub>2</sub> flux and transport models as a proxy for the background errors, spatial statistical tools were used to generate the background error statistics. The resultant error statistics were consistent with the large-scale structures in the background error, and implied errors correlated over longer distances relative to those deduced from the NMC statistics. This allowed the information from the assimilated observations to reduce errors over larger areas. For the test cases explored here, experiments using GOSAT CO<sub>2</sub> observations and subsequent evaluation with independent CO<sub>2</sub> observations illustrated that taking into account the errors in the background CO<sub>2</sub> fluxes is necessary to improve the assimilated atmospheric CO<sub>2</sub> concentrations. Because the variability in the CO<sub>2</sub> background, especially near the surface, is primarily driven by the underlying variability in the surface fluxes, this study makes the clear case for including a good estimate of surface flux errors and error correlations in any method that is used to estimate the background error statistics within a CO<sub>2</sub> data assimilation system.

[49] The experiments also demonstrated some caveats associated with the proposed  $\Delta$ -statistics approach. First, a judicious selection of the flux and transport models is recommended to capture realistic uncertainty in the background. As different combination of models may have different representations of the underlying mechanistic processes and/or atmospheric transport, and spatial and longitudinal biases may crop up near the surface due to inaccurate representations of these processes. Second, unlike the statistics from a forecast-based technique such as the NMC method, the generation of the  $\Delta$ -statistics requires additional analysis and computational time. Future work will investigate possible ways to generate the error statistics from an ensemble of models to reduce any form of biases and also reduce the overall computational time required to generate the background error statistics. The overall benefits from the new approach, however, outweigh these drawbacks, as demonstrated by a superior match to independent data relative to the existing NMC method. This exercise also provided a general insight into the improvements that can be achieved within operational atmospheric data assimilation systems from better representations of background error statistics.

[50] **Acknowledgments.** The authors thank Derek Posselt, Peter Adriaens and three anonymous reviewers for fruitful comments and discussions regarding this work, the many people at ECMWF who helped build the tracer data assimilation system, and partners within the GEMS and MACC projects. This work was supported by the National Aeronautics and Space Administration (NASA) through Earth System Science Fellowship for Abhishek Chatterjee, under grant NNX09AO10H. Additional support was provided through NASA grant NNX12AB90G. The work of Richard Engelen was funded through the MACC project, which is funded by the European Commission under the Seventh Research Framework Programme, contract 218793. The work of Stephan Kawa was supported through the NASA Carbon Cycle Science and Atmospheric CO<sub>2</sub> Observations from Space opportunities. TCCON data were obtained from the TCCON Data Archive, operated by the California Institute of Technology from the website at <http://tcon.ipac.caltech.edu/>. Finally, the GOSAT-ACOS data were produced by the ACOS/OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology, and obtained from the ACOS/OCO-2 data archive maintained at the NASA Goddard Earth Science Data and Information Services Center.

## References

- Alkhaled, A. A., A. M. Michalak, S. Randolph Kawa, S. C. Olsen, and J.-W. Wang (2008), A global evaluation of the regional spatial variability of column integrated CO<sub>2</sub> distributions, *J. Geophys. Res.*, **113**, D20303, doi:10.1029/2007JD009693.
- Andrews, A. E., K. A. Boering, S. C. Wofsy, B. C. Daube, D. B. Jones, S. Alex, M. Loewenstein, J. R. Podolske, and S. E. Strahan (2001), Empirical age spectra for the midlatitude lower stratosphere from in situ observations of CO<sub>2</sub>: Quantitative evidence for a subtropical “barrier” to horizontal transport, *J. Geophys. Res.*, **106**(D10), 10,257–10,274, doi:10.1029/2000JD900703.
- Aumann, H. H., et al. (2003), AIRS/AMSU/HSB on the aqua mission: Design, science objectives, data products, and processing systems, *IEEE Trans. Geosci. Remote Sens.*, **41**(2), doi:10.1109/tgrs.2002.808356.
- Baker, D. F., H. Bösch, S. C. Doney, D. O’Brien, and D. S. Schimel (2010), Carbon source/sink information provided by column CO<sub>2</sub> measurements from the Orbiting Carbon Observatory, *Atmos. Chem. Phys.*, **10**, 4145–4165, doi:10.5194/acp-10-4145-2010.
- Bannister, R. N. (2008a), A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, *Q. J. R. Meteorol. Soc.*, **134**(637), 1951–1970, doi:10.1002/qj.339.
- Bannister, R. N. (2008b), A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics, *Q. J. R. Meteorol. Soc.*, **134**(637), 1971–1996, doi:10.1002/qj.340.
- Belo Pereira, M., and L. Berre (2006), The use of an ensemble approach to study the background error covariances in a global NWP model, *Mon. Weather Rev.*, **134**(9), 2466–2489, doi:10.1175/mwr3189.1.
- Benedetti, A., and M. Fisher (2007), Background error statistics for aerosols, *Q. J. R. Meteorol. Soc.*, **133**(623), 391–405, doi:10.1002/qj.37.
- Berre, L., and G. Desroziers (2010), Filtering of background error variances and correlations by local spatial averaging: A review, *Mon. Weather Rev.*, **138**(10), 3693–3720, doi:10.1175/2010mwr3111.1.
- Berre, L., S. E. Stefanescu, and M. B. Pereira (2006), The representation of the analysis effect in three error simulation techniques, *Tellus Ser. A. Dyn. Meteorol. Oceanogr.*, **58**(2), 196–209, doi:10.1111/j.1600-0870.2006.00165.x.
- Bonavita, M., L. Raynaud, and L. Isaksen (2011), Estimating background-error variances with the ECMWF Ensemble of Data Assimilations system: Some effects of ensemble size and day-to-day variability, *Q. J. R. Meteorol. Soc.*, **137**(655), 423–434, doi:10.1002/qj.756.
- Bönsch, H., A. Engel, J. Curtius, T. Birner, and P. Hoor (2009), Quantifying transport into the lowermost stratosphere using simultaneous in-situ measurements of SF<sub>6</sub> and CO<sub>2</sub>, *Atmos. Chem. Phys.*, **9**(16), 5905–5919, doi:10.5194/acp-9-5905-2009.
- Bösch, H., et al. (2006), Space-based near-infrared CO<sub>2</sub> measurements: Testing the orbiting carbon observatory retrieval algorithm and validation concept using SCIAMACHY observations over Park Falls, Wisconsin, *J. Geophys. Res.*, **111**, D23302, doi:10.1029/2006JD007080.
- Brenninkmeijer, C. A. M., et al. (2007), Civil aircraft for the regular investigation of the atmosphere based on an instrumented container: The new CARIBIC system, *Atmos. Chem. Phys.*, **7**(18), 4953–4976, doi:10.5194/acp-7-4953-2007.
- Brousseau, P., L. Berre, F. Bouttier, and G. Desroziers (2012), Flow-dependent background-error covariances for a convective-scale data assimilation system, *Q. J. R. Meteorol. Soc.*, **138**(663), 310–322, doi:10.1002/qj.920.
- Buehner, M., and M. Charron (2007), Spectral and spatial localization of background-error correlations for data assimilation, *Q. J. R. Meteorol. Soc.*, **133**(624), 615–630, doi:10.1002/qj.50.



- Cardinali, C., S. Pezzulli, and E. Andersson (2004), Influence-matrix diagnostic of a data assimilation system, *Q. J. R. Meteorol. Soc.*, **130**(603), 2767–2786, doi:10.1256/qj.03.205.
- Chai, T. F., et al. (2007), Four-dimensional data assimilation experiments with International Consortium for Atmospheric Research on Transport and Transformation ozone measurements, *J. Geophys. Res.*, **112**, D12S15, doi:10.1029/2006JD007763.
- Chevallier, F., R. J. Engelen, C. Carouge, T. J. Conway, P. Peylin, C. Pickett-Heaps, M. Ramonet, P. J. Rayner, and I. Xueref-Remy (2009a), AIRS-based versus flask estimation of carbon surface fluxes, *J. Geophys. Res.*, **114**, D20303, doi:10.1029/2009JD012311.
- Chevallier, F., S. Maksyutov, P. Bousquet, F.-M. Bréon, R. Saito, Y. Yoshida, and T. Yokota (2009b), On the accuracy of the CO<sub>2</sub> surface fluxes to be estimated from the GOSAT observations, *Geophys. Res. Lett.*, **36**, L19807, doi:10.1029/2009GL040108.
- Chiles, J.-P., and P. Delfiner (2012), *Geostatistics: Modeling Spatial Uncertainty* (Wiley Series in Probability and Statistics), 734 pp., John Wiley & Sons, Inc., Hoboken, New Jersey.
- Constantinescu, E. M., T. Chai, A. Sandu, and G. R. Carmichael (2007), Autoregressive models of background errors for chemical data assimilation, *J. Geophys. Res.*, **112**, D12309, doi:10.1029/2006JD008103.
- Crevoisier, C., A. Chedin, H. Matsueda, T. Machida, R. Armante, and N. A. Scott (2009), First year of upper tropospheric integrated content of CO<sub>2</sub> from IASI hyperspectral infrared observations, *Atmos. Chem. Phys.*, **9**(14), 4797–4810, doi:10.5194/acp-9-4797-2009.
- Crevoisier, C., C. Sweeney, M. Gloor, J. L. Sarmiento, and P. P. Tans (2010), Regional US carbon sinks from three-dimensional atmospheric CO<sub>2</sub> sampling, *Proc. Natl. Acad. Sci. U. S. A.*, **107**(43), 18,348–18,353, doi:10.1073/pnas.0900062107.
- Crisp, D., et al. (2012), The ACOS CO<sub>2</sub> retrieval algorithm. Part II: Global XCO<sub>2</sub> data characterization, *Atmos. Meas. Tech.*, **5**, 687–707, doi:10.5194/amt-5-687-2012.
- Denning, A. S., T. Takahashi, and P. Friedlingstein (1999), Can a strong atmospheric CO<sub>2</sub> rectifier effect be reconciled with a “reasonable” carbon budget?, *Tellus Ser. B-Chem. Phys. Meteorol.*, **51**(2), doi:10.1034/j.1600-0889.1999.t01-1-00010.x.
- Derber, J., and F. Bouttier (1999), A reformulation of the background error covariance in the ECMWF global data assimilation system, *Tellus Ser. A-Dyn. Meteorol. Oceanogr.*, **51**(2), 195–221, doi:10.1034/j.1600-0870.1999.t01-2-00003.x.
- Engelen, R. J., and A. P. McNally (2005), Estimating atmospheric CO<sub>2</sub> from advanced infrared satellite radiances within an operational four-dimensional variational (4D-Var) data assimilation system: Results and validation, *J. Geophys. Res.*, **110**, D18305, doi:10.1029/2005JD005982.
- Engelen, R. J., E. Andersson, F. Chevallier, A. Hollingsworth, M. Matricardi, A. P. McNally, J.-N. Thepaut, and P. D. Watts (2004), Estimating atmospheric CO<sub>2</sub> from advanced infrared satellite radiances within an operational 4D-Var data assimilation system: Methodology and first results, *J. Geophys. Res.*, **109**, D19309, doi:10.1029/2004JD004777.
- Engelen, R. J., S. Serrar, and F. Chevallier (2009), Four-dimensional data assimilation of atmospheric CO<sub>2</sub> using AIRS observations, *J. Geophys. Res.*, **114**, D03303, doi:10.1029/2008JD010739.
- Fisher, M. (2004), Background error covariance modelling, Proc. of ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean, Reading, UK, 45–64.
- Fisher, M. (2006), “Wavelet”Jb –A new way to model the statistics of background errors. ECMWF Newsletter, **106**, 23–28, *Society*, **130**(601), 2253–2275, doi:10.1256/qj.03.26.
- Gurk, C., H. Fischer, P. Hoor, M. G. Lawrence, J. Lelieveld, and H. Wernli (2008), Airborne in-situ measurements of vertical, seasonal and latitudinal distributions of carbon dioxide over Europe, *Atmos. Chem. Phys.*, **8**(21), 6395–6403, doi:10.5194/acp-8-6395-2008.
- Hammerling, D. M., A. M. Michalak, and S. R. Kawa (2012a), Mapping of CO<sub>2</sub> at high spatiotemporal resolution using satellite observations: Global distributions from OCO-2, *J. Geophys. Res.*, **117**, D06306, doi:10.1029/2011JD017015.
- Hammerling, D. M., A. M. Michalak, C. O'Dell, and S. R. Kawa (2012b), Global CO<sub>2</sub> distributions over land from the Greenhouse Gases Observing Satellite (GOSAT), *Geophys. Res. Lett.*, **39**, L08804, doi:10.1029/2012GL051203.
- Hess, R. (2010), Flow dependence of background errors and their vertical correlations for radiance-data assimilation, *Q. J. R. Meteorol. Soc.*, **136**(647), 475–486, doi:10.1002/qj.570.
- Hollingsworth, A., et al. (2008), Toward a monitoring and forecasting system for atmospheric composition: The GEMS project, *Bull. Am. Meteorol. Soc.*, **89**(8), 1147–1164, doi:10.1175/2008bams2355.1.
- Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglass, R. B. Rood, and L. Pfister (1995), Stratosphere-troposphere exchange, *Rev. Geophys.*, **33**(4), 403–439, doi:10.1029/95RG02097.
- Huntzinger, D. N., et al. (2012), North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison, *Ecol. Model.*, **232**, doi:10.1016/j.ecolmodel.2012.02.004.
- Kahnert, M. (2008), Variational data analysis of aerosol species in a regional CTM: Background error covariance constraint and aerosol optical observation operators, *Tellus Ser. B-Chem. Phys. Meteorol.*, **60**(5), 753–770, doi:10.1111/j.1600-0889.2008.00377.x.
- Kawa, S. R., D. J. Erickson, S. Pawson, and Z. Zhu (2004), Global CO<sub>2</sub> transport simulations using meteorological data from the NASA data assimilation system, *J. Geophys. Res.*, **109**, D18312, doi:10.1029/2004JD004554.
- Krinner, G., et al. (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochem. Cycles*, **19**, GB1015, doi:10.1029/2003GB002199.
- Kulawik, S. S., N. Viovy, N. de Noblet-Ducoudre, J. Ogée, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch, and I. Colin Prentice (2010), Characterization of Tropospheric Emission Spectrometer (TES) CO<sub>2</sub> for carbon cycle science, *Atmos. Chem. Phys.*, **10**(12), 5601–5623, doi:10.5194/acp-10-5601-2010.
- Kuze, A., H. Suto, M. Nakajima, and T. Hamazaki (2009), Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, *Appl. Opt.*, **48**, 6716–6733, doi:10.1364/AO.48.006716.
- Le Quere, C., et al. (2009), Trends in the sources and sinks of carbon dioxide, *Nat. Geosci.*, **2**(12), 831–836, doi:10.1038/ngeo689.
- Liu, J., I. Fung, E. Kalnay, J.-S. Kang, E. T. Olsen, and L. Chen (2012), Simultaneous assimilation of AIRS XCO<sub>2</sub> and meteorological observations in a carbon climate model with an ensemble Kalman filter, *J. Geophys. Res.*, **117**, D05309, doi:10.1029/2011JD016642.
- Machida, T., N. Vivo, N. de Noblet-Ducoudre, J. Ogée, J. Polcher, P. Friedlingstein, P. Ciais, S. Stephen, and I. C. Prentice (2008), Worldwide measurements of atmospheric CO<sub>2</sub> and other trace gas species using commercial airlines, *J. Atmos. Oceanic Technol.*, **25**(10), 1744–1754, doi:10.1175/2008JTECHA1082.1.
- Massart, S., A. Piacentini, and O. Pannekoucke (2012), Importance of using ensemble estimated background error covariances for the quality of atmospheric ozone analyses, *Q. J. R. Meteorol. Soc.*, **138**(665), 889–905, doi:10.1002/qj.971.
- Nichols, D. (2010), Mathematical concepts of data assimilation, in *Part I. Theory, Data Assimilation, Making Sense of Observations*, edited by W. Lahoz, B. Khattatov, and R. Menard, pp. 13–40, Springer-Verlag, Berlin.
- Nocedal, J., and S. J. Wright (2006), *Numerical Optimization*, Springer Ser. Oper. Res., pp. 224–229, Springer-Verlag, Berlin.
- O'Dell, C. W., et al. (2012), The ACOS CO<sub>2</sub> retrieval algorithm. Part 1: Description and validation against synthetic observations, *Atmos. Meas. Tech.*, **5**(1), 99–121, doi:10.5194/amt-5-99-2012.
- Pannekoucke, O., L. Berre, and G. Desroziers (2007), Filtering properties of wavelets for local background-error correlations, *Q. J. R. Meteorol. Soc.*, **133**(623), 363–379, doi:10.1002/qj.33.
- Pannekoucke, O., L. Berre, and G. Desroziers (2008), Background-error correlation length-scale estimates and their sampling statistics, *Q. J. R. Meteorol. Soc.*, **134**(631), 497–508, doi:10.1002/qj.212.
- Parrish, D. F., and J. C. Derber (1992), The National Meteorological Centre's spectral statistical-interpolation analysis system, *Mon. Weather Rev.*, **120**(8), 1747–1763, doi:10.1175/1520-0493(1992)120<1747:tmcss>2.0.co;2.
- Randerson, J. T., M. V. Thomps, T. J. Conw, I. Y. Fun, and C. B. Field (1997), The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, *Global Biogeochem. Cycles*, **11**(4), 535–560, doi:10.1029/97GB02268.
- Raynaud, L., L. Berre, and G. Desroziers (2009), Objective filtering of ensemble-based background-error variances, *Q. J. R. Meteorol. Soc.*, **135**(642), 1177–1199, doi:10.1002/qj.438.
- Scholes, R. J., P. M. S. Monteiro, C. L. Sabine, and J. G. Canadell (2009), Systematic long-term observations of the global carbon cycle, *Trends Ecol. Evol.*, **24**(8), 427–430, doi:10.1016/j.tree.2009.03.006.
- Simeoni, D., C. Singer, and G. Chalon (1997), Infrared atmospheric sounding interferometer, *Acta Astronaut.*, **40**(2–8), doi:10.1016/s0094-5765(97)00098-2.
- Singh, K., M. Jarak, A. Sandu, K. Bowman, M. Lee, and D. Jones (2011), Construction of non-diagonal background error covariance matrices for global chemical data assimilation, *Geoscientific Model Dev.*, **4**(2), 299–316, doi:10.5194/gmd-4-299-2011.
- Siroka, M., C. Fischer, V. Casse, R. Brozkova, and J. F. Geleyn (2003), The definition of mesoscale selective forecast error covariances for a limited area variational analysis, *Meteorol. Atmos. Phys.*, **82**(1–4), 227–244, doi:10.1007/s00703-001-0588-5.
- Stohl, A., S. Eckhardt, C. Forster, P. James, and N. Spichtinger (2002), On the pathways and timescales of intercontinental air pollution transport, *J. Geophys. Res.*, **107**(D23), 4684, doi:10.1029/2001JD001396.



- Storto, A., and R. Randriamampianina (2010), Ensemble variational assimilation for the representation of background error covariances in a high-latitude regional model, *J. Geophys. Res.*, *115*, D17204, doi:10.1029/2009JD013111.
- Tans, P. P. (1996), Carbon cycle (group report), in *Summary Report 1994–1995*, vol. 23, edited by D. J. Hoffman, J. Peterson, and R. M. Rosson, pp. 29–49, U. S. Dep. of Commer, Boulder, Colorado.
- Tiwari, Y. K., M. Gloor, R. J. Engelen, F. Chevallier, C. Rödenbeck, S. Körner, P. Peylin, B. H. Braswell, and M. Heimann (2006), Comparing CO<sub>2</sub> retrieved from Atmospheric Infrared Sounder with model predictions: Implications for constraining surface fluxes and lower-to-upper troposphere transport, *J. Geophys. Res.*, *111*, D17106, doi:10.1029/2005JD006681.
- Willmott, C. J., and K. Matsuura (2005), Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, *Clim. Res.*, *30*(1), 79–82, doi:10.3354/cr030079.
- Wofsy, S. C. (2011), HIPER Pole-to-Pole Observations (HIPPO): Fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, *Proc. R. Soc. A*, *369*, 2073–2086, doi:10.1098/rsta.2010.0313.
- Wunch, D., et al. (2011a), A method for evaluating bias in global measurements of CO<sub>2</sub> total columns from space, *Atmos. Chem. Phys.*, *11*, 12,317–12,337, doi:10.5194/acp-11-12317-2011.
- Wunch, D., G. C. Toon, J.-F. L. Blavier, R. A. Washenfelder, J. Notholt, B. J. Connor, D. W. T. Griffith, V. Sherlock, and P. O. Wennberg (2011b), The Total Carbon Column Observing Network, *Phil. Trans. R. Soc. A-Math. Phys. Eng. Sci.*, *369*(1943), 2087–2112, doi:10.1098/rsta.2010.0240.
- Xueref-Remy, I., P. Bousquet, C. Carouge, L. Rivier, and P. Ciais (2011), Variability and budget of CO<sub>2</sub> in Europe: Analysis of the CAATER airborne campaigns. Part 2: Comparison of CO<sub>2</sub> vertical variability and fluxes between observations and a modeling framework, *Atmos. Chem. Phys.*, *11*(12), doi:10.5194/acp-11-5673-2011.
- Yokota, T., Y. Yoshida, N. Eguchi, Y. Ota, T. Tanaka, H. Watanabe, and S. Maksyutov (2009), Global concentrations of CO<sub>2</sub> and CH<sub>4</sub> retrieved from GOSAT, *SOLA*, *5*, 160–163, doi:10.2151/sola.2009-041.